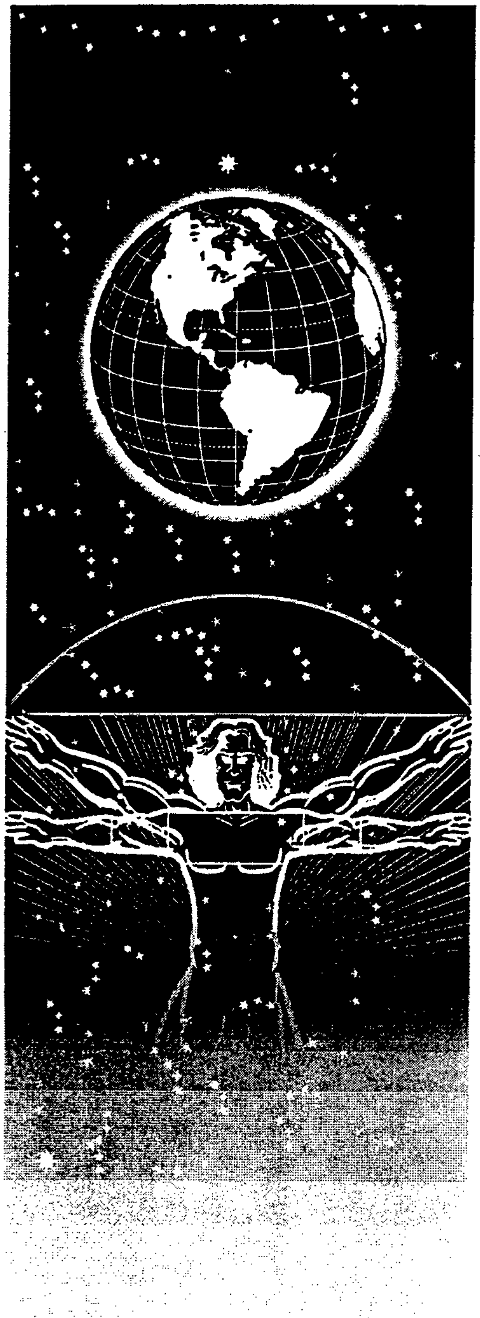
AFRL-HE-AZ-TR-1998-0076



UNITED STATES AIR FORCE RESEARCH LABORATORY

FACTOR STRUCTURE OF THE COGSCREEN-AERONAUTICAL EDITION TEST BATTERY

Thomas R. Carretta

Malcolm James Ree

Training Effectiveness Branch
Warfighter Training Research Division
7909 Lindbergh Drive
Brooks Air Force Base TX 78235-5352

Joseph D. Callister

USAF School of Aerospace Medicine
Aeromedical Consult Service
2507 Kennedy Circle
Brooks Air Force Base TX 78235-5117

May 1999

AIR FORCE MATERIEL COMMAND
AIR FORCE RESEARCH LABORATORY
HUMAN EFFECTIVENESS DIRECTORATE
WARFIGHTER TRAINING RESEARCH DIVISION
6030 South Kent
Mesa AZ 85212-6004

NOTICES

Using Government drawings, specifications, or other data included in this document for any purpose other than Government-related procurement does not in any way obligate the US Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation, or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

THOMAS R. CARRETTA
Project Scientist

DEE H. ANDREWS
Technical Director

JERALD L. STRAW, Colonel, USAF
Chief, Warfighter Training Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
May 1999

3. REPORT TYPE AND DATES COVERED
Final - March 1995 to October 1997

4. TITLE AND SUBTITLE

Factor Structure of the CogScreen-Aeronautical Edition Test Battery

5. FUNDING NUMBERS

PE - 62202F
PR - 1123
TA - B1
WU - 01

6. AUTHOR(S)

Carretta, Thomas R., Ree, Malcolm James, & *Callister, J.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory, Human Effectiveness Directorate
Warfighter Training Research Division, Training Effectiveness Branch
7909 Lindbergh Drive; Brooks Air Force Base TX 78235-5352
*USAF School of Aerospace Medicine
2507 Kennedy Circle; Brooks AFB TX 78235-5117

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6030 South Kent
Mesa AZ 85212-6004

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AFRL-HE-AZ-TR-1998-0076

11. SUPPLEMENTARY NOTES

Air Force Research Laboratory Technical Monitor: Dr Thomas R. Carretta, (937) 656-7014; DSN 986-7014

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

CogScreen-Aeronautical Edition (CogScreen-AE) is a computerized test designed to assess cognitive functioning. The US Air Force (USAF) is evaluating its utility for establishing a cognitive baseline for pilots that would assist clinicians when evaluating pilots with cognitive referral, which could result in impaired flying ability. To better understand what is measured by the CogScreen-AE, confirmatory factor analyses were performed. Participants were 1,015 USAF pilot training applicants. Like many cognitive tests, CogScreen-AE exhibited an hierarchical factor structure. Somewhat unusual was that there was not a strong, single, higher order factor representing general cognitive ability. Instead, there were two higher order factors that measured response efficiency and procedural knowledge. In addition to the higher order factors, there were six first-order factors representing thruput, response speed, shifting attention, psychomotor tracking, pathfinder, and numeric.

14. SUBJECT TERMS
Cognitive ability; Cognitive functioning; Cogscreen-Aeronautical Edition; Cogscreen-AE; Pilot selection; Pilot training; Test battery; Tests;

15. NUMBER OF PAGES
21

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION ABSTRACT
Unclassified	Unclassified	Unclassified	UL

CONTENTS

FIGURES

TABLES

PREFACE

This effort was conducted under Work Unit 1123-B1-01, Pilot Selection and Classification Support, in support of aircrew selection and classification research and development at the Air Force Research Laboratory, Human Effectiveness Directorate, Warfighter Training Research Division (AFRL/HEA), Brooks Air Force Base, TX. The laboratory work unit monitor was Dr Thomas R. Carretta. Dr Carretta is currently associated with AFRL/HECI at Wright-Patterson Air Force Base, OH. He may be reached via e-mail at thomas.carretta@he.wpafb.af.mil.

We would like to thank SSgt Pauline Etterle for her assistance in reviewing the CogScreen-AE battery, and thank Dr Andrea Berndt for her help in specifying the LISREL models.

FACTOR STRUCTURE OF THE COGSCREEN-AE TEST BATTERY

SUMMARY

CogScreen Aeronautical Edition (CogScreen-AE) is a computerized test designed to assess cognitive functioning. The US Air Force (USAF) is evaluating its utility for establishing a cognitive baseline for pilots that would assist clinicians when evaluating pilots with cognitive referral, which could result in impaired flying ability. To better understand what is measured by the CogScreen-AE, confirmatory factor analyses were performed. Participants were 1,015 USAF pilot training applicants. Like many cognitive tests, CogScreen-AE exhibited an hierarchical factor structure. Somewhat unusual was that there was not a strong, single, higher order factor representing general cognitive ability. Instead, there were two higher order factors that measured response efficiency and procedural knowledge. In addition to the higher order factors, there were six first-order factors representing thruput, response speed, shifting attention, psychomotor tracking, pathfinder, and numeric.

INTRODUCTION

Ability tests have long been used in the selection of pilots, especially in the military (Hilton & Dolgin, 1991; Ree & Carretta, 1996). These tests have been both cognitive and psychomotor and have been administered via paper-and-pencil and by computer (Ree & Carretta, 1998). Another use for ability tests is to detect minimal brain dysfunction. The US Federal Aviation Administration (FAA) called for such a test in 1987. This study investigated the factor structure of a computerized test designed for detecting deficits in the cognitive functioning of pilots due to minimal brain dysfunction.

CogScreen-Aeromedical Edition (CogScreen-AE) (Kay, 1995) is a computer-based cognitive ability test intended to assess changes or deficits in cognitive, information processing, and perceptual abilities (e.g., attention, memory, reaction time, simultaneous information processing). It was initially designed to detect subtle changes in cognitive functioning of aviation pilots that could result in impaired flying ability (Kay, 1995). Unlike other tests of aviation aptitude, CogScreen-AE does not measure aviation knowledge (e.g., Air Force Officer Qualifying Test or AFOQT) (Carretta & Ree, 1996) or flying skills (e.g., Canadian Automated Pilot Selection System or CAPSS) (Spinner, 1991). Rather, it focuses on the measurement of cognitive, information processing, and perceptual abilities hypothesized to be related to flying performance (Imhoff & Levine, 1981). The US Air Force (USAF) is evaluating the utility of the CogScreen-AE battery along with a computerized version of the Multidimensional Aptitude Battery (MAB) (Jackson, Barton, & Blokker, 1992) for establishing a cognitive baseline for student pilots (Callister, King, & Retzlaff, 1995; Retzlaff, Callister, & King, 1996).

Kay (1995) conducted exploratory factor analyses of 28 CogScreen-AE scores for an international sample of 662 United States and Russian pilots. Principal components analysis with Varimax rotation was used. A 9-factor solution was generated that accounted for 67.1% of the

variance. The factors were: (1) visual scanning and sequencing, (2) attribute identification, (3) visual perceptual and spatial processing, (4) motor coordination, (5) choice visual reaction time, (6) visual associative memory, (7) tracking, (8) working memory, and (9) numerical operations. The factors were not allowed to correlate, as an orthogonal solution was specified. However, interpretation of these factors is questionable. Four of the nine factors were underspecified (i.e., represented by two or fewer scores). Finally, Kay did not test the statistical goodness-of-fit of his factor model in his exploratory factor analysis.

The purposes of this study were to test the goodness-of-fit of Kay's factor model and to establish and test the goodness-of-fit of two alternate factor models of the CogScreen-AE as established by cumulative psychometric research on ability tests (Carretta & Ree, 1996, 1997; Carretta, Retzlaff, Callister, & King, 1998; Jensen, 1980; Ree & Carretta, 1994; Stauffer, Ree, & Carretta, 1996; Vernon, 1969).

METHOD

Participants

Participants were 1,015 US Air Force pilot training applicants. The mean age at time of testing was 20.5 years. The sample was predominantly male (91.9%) and White (88.4%). Participants were tested as a requirement for entrance into the US Air Force Enhanced Flight Screening Program. This program teaches basic flying skills and is used to determine which applicants should attend jet aircraft training.

The protocol for the current study had been reviewed and approved by the Air Force Human Use Committee of the Air Force Medical Operations Agency. Informed consent was obtained from participants before their participation.

Measures

The CogScreen-AE tests were administered on a 386-based computer with a 14-inch color monitor. Participants entered their responses using a keypad and mouse or light pen. Table 1 provides the test developer's brief summary of the attributes measured by the CogScreen-AE tests and the types of scores generated for each test. Most of the tests include percent correct, median response time for items answered correctly, and thruput. Thruput is defined as a measure of response efficiency and reflects the number of correct responses per minute (Kay, 1995, p.8). Some CogScreen-AE tests (i.e., Divided Attention, Pathfinder, Selective Attention, and Dual Task) include various process-oriented, qualitative, or response error measures that cannot be classified as accuracy, speed, or thruput. Brief descriptions of the tests follow. Kay (1995) provided detailed test descriptions and scoring procedures.

Backward digit span. This test is similar to the backward digit span task found in the Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler, 1981). In this test, participants are sequentially presented with a series of 3 to 6 digits. Their task is to reproduce the sequence in reverse order. The score is the percent of series recalled in correct reverse order (BDSACC).

Math. This test consists of traditional multistep word problems. Scores include percent correct (MATHACC), median response time on items answered correctly (MATHRTC), and thruput (i.e., MATHPUT, number of problems answered correctly per minute).

Table 1. Brief Descriptions of CogScreen-AE Tests

Test	Abbreviation	Attributes Measured	Types of Scores
Backward Digit Span	BDS	verbal sequential processing, visual attention, & working memory	response accuracy
Math	MATH	math skills, reading comprehension, reasoning, & working memory	response accuracy, response speed, & thruput
Visual Sequence Comparison	VSC	verbal sequential processing, visual attention, visual perceptual speed, & working memory	response accuracy, response speed, & thruput
Symbol Digit Coding	SDC	attention, immediate and delayed memory, information processing	response accuracy, response speed, & thruput
Matching-to-Sample	MTS	spatial processing, visual-perceptual speed, & working memory	response accuracy, response speed, & thruput
Manikin	MAN	spatial orientation & visual-spatial perception	response accuracy, response speed, & thruput
Divided Attention	DAT	choice reaction time, divided attention, multitasking, & working memory	response accuracy, response speed, & thruput
Auditory Sequence Comparison	ASC	auditory attention, sound pattern discrimination, & working memory	response accuracy, response speed, & thruput
Pathfinder	PF	ability to shift mental set, memory, motor coordination, organizing and sequencing, & visual scanning	response accuracy, response speed, & thruput
Shifting Attention	SAT	application of rules, choice reaction time, concept formation, deductive reasoning, & working memory	response accuracy, response speed, & thruput
Dual Task	DTT	divided attention, multitasking, & working memory	response accuracy, response speed, thruput, & tracking error

Visual sequence comparison. Two alphanumeric strings are simultaneously presented side-by-side on the screen. Participants must determine whether the strings are the same or different. String length varies from 4 to 8 characters. Performance measures include percent correct (VSCACC), response time for correctly answered items (VSCRTC), and thruput (VSCPUT).

Symbol digit coding. This test consists of three tasks: (1) symbol-digit coding, (2) immediate recall, and (3) delayed recall. The symbol-digit coding task is a computer-based analogue of Digit Symbol test from the WAIS-R (Wechsler, 1981). In this task, six symbol-digit pairs are shown at the top of the screen. The pairs remain on the screen throughout the symbol-digit-coding task. Participants are instructed to remember the symbol-digit pairs for a later memory test (immediate and delayed recall tasks; see below). Meanwhile, in the center of the screen, a row of symbols is shown with blank spaces below each symbol. Participants enter the digit that corresponds to each symbol. Scores include accuracy (SDCACC), and the number of items completed correctly per minute (i.e., thruput, SDCPUT).

The immediate recall task measures the ability to recall the symbol-digit pairs presented in the first part of the test. The six symbols used in the symbol-digit-coding task are presented. Participants are required to remember the corresponding digit. Response accuracy (SDCIRACC) is the only score.

The delayed recall task is identical to the immediate recall task, except that it is presented after the Pathfinder test. As with the immediate recall task, the only score is response accuracy (SDCDRACC).

Matching-to-sample. Participants are shown a 4 x 4 grid of filled and empty cells (i.e., target grid). After a brief presentation, the target grid is removed and replaced by two grids that are displayed side-by-side on the screen. One pattern is identical to the target grid while the other differs in 1 of the 16 cells. Participants must choose the grid matching the target grid. Scores include accuracy (MTSACC), response time (MTSRTC), and thruput (MTSPUT).

Manikin. This test is based on a spatial transformation task described by Benson and Gedye (1963). Participants are presented with a male figure in one of four positions. The image can be either right side up or upside down and facing toward or facing away. The figure holds a flag in one hand. Participants must determine which hand (right or left) is holding the flag. Test performance is scored for accuracy (MANACC), speed (MANRTC), and thruput (MANPUT).

Divided attention test. This test has two parts: (1) a visual indicator monitoring task and (2) the visual indicator monitoring task presented simultaneously with the Visual Sequence Comparison task (see above). In the first part (visual monitoring alone task), participants watch a cursor move vertically within a circle divided into upper, central, and lower sections. When the cursor moves from the central section into either the upper or lower section, participants touch a box labeled "CENTER" with a light pen to recenter the cursor. Scores include the median amount of time the cursor spends in either the upper or lower sections before being recentered (DATIRTC) and the number of premature (i.e., anticipatory) responses (DATIPRE).

In part 2 (dual-task), the visual indicator-monitoring task is paired with the Visual Sequence Comparison task. Performance measures include response speed (DATDRTC) and number of premature responses (DATDPRE) for the indicator monitoring task and response accuracy (DATSCACC), response speed (DATSCRTC), and thruput (DATSCPUT) for the visual sequence comparison task.

Auditory sequence comparison. This test involves the comparison of two series of 4 to 8 tones. The series are presented sequentially and may be identical or differ in the pitch of one tone. Participants make a "SAME" or "DIFFERENT' judgment for each pair. Scores include accuracy (ASCACC), response speed (ASCRTC), and thruput (ASCPUT).

Pathfinder. Participants are taught three sequencing rules regarding numbers, letters, and a combination of numbers and letters. For each item, a number or letter is shown in the center of the screen. Participants choose from among four characters (numbers and/or letters), the one that should appear next in the sequence. Three of the four character choices are updated following each response.

There are separate accuracy, speed, thruput, and coordination scores for letter, number, and "combination" items. The response accuracy scores are PFLACC (letters), PFNACC (numbers), and PFCACC (combined). The median response time scores are PFLMRT (letters), PFNMRT (numbers), and PFCMRT (combined). The thruput scores are PFLPUT (letters), PFNPUT (numbers), and PFCPUT (combined). Coordination scores measure participants proximity to the center of the target numbers and letters and include PFLCOOR (letters), PFNCOOR (numbers), and PFCCOOR (combined).

Shifting attention test. In this test, participants must alter their responses depending on rule changes. Under Rule 1, boxes are selected based on the color of their borders. Under Rule 2, boxes are selected based on the direction of their arrow, and under Rule 3, they are selected based on the color of their arrow. The test items begin after the three rules have been presented and a practice session completed. During the first group of test items, one of the rules is shown before each item. In the second group of test items (discovery), participants must discover, and then apply, the active response rule, which changes after a variable number of correct answers. During the discovery items, participants must use trial-and-error to determine the active rule.

Response accuracy, speed, and thruput are measured for all tasks in this test. Additional scores for the discovery items include number of shifts completed (SATDIRUL), failures to maintain set (SATDIFAI), number of nonconceptual responses (SATDINON), and number of perseverative errors (SATDIPER).

Dual task. This test is similar to various measures of time-sharing ability (see, for example, Carretta & Ree, 1993; North & Gopher, 1976). In this test, participants must learn to separately perform a tracking task and a memory task, then perform both simultaneously. In the first part (compensatory tracking alone), participants use the left and right arrow keys to keep a cursor in the center of the screen. Without intervention, the cursor will drift off the screen causing a "boundary hit." Scores for part 1 include mean absolute tracking error (DTTAABS) and the number of boundary hits (DTTAHIT).

The second part involves a delayed recall task. In this task, participants are shown the numbers 1, 2, or 3. The first number is then replaced by a 1, 2, or 3. Participants must recall the previous number shown and select it (using a light pen), while simultaneously mentally encoding

the current number for the next item presentation. Performance measures include accuracy (DTTPAACC), response speed (DTTPARTC), and thruput (DTTPAPUT). The third part simultaneously presents the tracking and delayed recall tasks. Scores include tracking error (DTTDABS), boundary hits (DTTDHIT), recall accuracy (DTTPDACC), response speed (DTTPDRTC), and thruput (DTTPDPUT).

Procedures

Participants completed the CogScreen-AE tests shortly before beginning the USAF Enhanced Flight Screening Program. CogScreen-AE testing was done to establish an ideographic cognitive baseline for the clinical evaluation of pilots for comparative purposes after sustaining a head injury or other neurological insult (Retzlaff, et al., 1996).

Analyses included descriptive statistics, correlations, and confirmatory factor analyses. Because the participants had been selected, at least in part, on the basis of aptitude test scores, they constituted a range-restricted sample. Such samples provide relatively poor statistical estimates of the relations between variables (Thorndike, 1949). The multivariate correction method (Lawley, 1943; Ree, Carretta, Earles, & Albert, 1994) was used to correct the means, standard deviations, and correlations for the effects of prior selection to a US Air Force applicant sample (Skinner & Ree, 1987). The corrected correlation matrix was used in the confirmatory factor analyses.

Kay (1995) lists 65 summary scores from the CogScreen-AE tests (19 accuracy, 19 speed, 16 thruput, and 11 process-oriented). As noted earlier, the thruput scores are linear transformations of the accuracy and speed scores. Therefore, it would be inappropriate to include accuracy and speed scores in confirmatory factor analyses with thruput scores based on the same test items. We chose to follow Kay's lead by focusing on the same 28 scores as used in his previous exploratory factor analysis (3 accuracy, 5 speed, 12 thruput, and 8 process scores). Several models were specified, parameters estimated, and the goodness-of-fit to the data was tested. Hierarchical confirmatory factor analyses were performed using LISREL 8 (Jöreskog & Sörbom, 1996).

Model 1 consisted of the nine first-order orthogonal factors described by Kay (1995). Model 2 was identical to Model 1 except that the first-order factors were allowed to correlate as is usually found in cognitive ability tests. Model 3 was a simple, more parsimonious model with six first-order factors. Some of these factors were developed by grouping similar types of scores across tests (i.e., thruput, response time, tracking, numerical), while others represented specific CogScreen-AE tests with multiple scores per test (i.e., Shifting Attention has 7 scores, Pathfinder has 6 scores, etc.). The first-order confirmatory factor analyses (CFAs) allowed all observed variables (28 CogScreen-AE scores) to load on their first-order factors. This retains the factor influence on all scores and was necessary to conduct the latter hierarchical CFAs based on the correlations of the six first-order factors.

Two goodness-of-fit indices were computed to test the goodness of the models. These were the Comparative Fit Index (CFI) (Bentler, 1990) and Root Mean Square Error of

Approximation (RMSEA) (Browne & Cudeck, 1993). Values above .90 for the CFI and below .08 for the RMSEA are considered indicators of acceptable fit.

RESULTS AND DISCUSSION

The sample data were corrected for range restriction. The means and standard deviations are presented in Table 2 and the correlation matrix is presented in Table 3. As would be expected in selected samples, after correction for range restriction, the means decreased and the standard deviations increased. In general, these changes were small. After correction for range restriction, the correlations increased, but again by only a small amount. Because the Air Force participants were selected on highly *g*-loaded tests, this suggests that the CogScreen-AE is not heavily *g*-loaded. If it were, the restriction caused by incidental selection would have been greater (Thorndike, 1949) leading to larger differences between pre- and post-correction values. All confirmatory factor models were estimated using the corrected data.

Table 2. Means and Standard Deviations

Score	Abbr.	Mean	SD	Mean	SD
Backward Digit Span					
1. Accuracy	BDSACC	0.85	0.17	0.81	0.18
Math					
2. Thruput	MATHPUT	2.19	0.98	1.86	1.09
Visual Sequence Comparison					
3. Thruput	VCSPUT	29.84	6.65	29.10	6.77
Symbol Digit Coding					
4. Thruput	SDCPUT	76.63	23.62	72.04	24.17
5. Immediate Recall Acc.	SDCIRACC	0.12	0.31	0.11	0.31
Matching-to-Sample					
6. Thruput	MTSPUT	49.47	10.26	47.53	10.29
Manikin					
7. Thruput	MANPUT	34.27	8.94	30.97	9.19
Divided Attention Test					
8. Sequence Comparison Thruput	DATSCPUT	27.43	7.01	26.69	7.11
9. Indicator "Alone" RT	DATIRTC	0.37	0.09	0.37	0.09
10. Indicator "Dual" RT	DATDRTC	0.63	0.19	0.63	0.20

Table 2. Means and Standard Deviations (Cont'd)

Score	Abbr.	Mean	SD	Mean	SD
Auditory Sequence Comp.					
11. Thruput	ASCPUT	92.90	27.10	90.25	27.25
Pathfinder					
12. Letter RT	PFLRTC	0.70	0.18	0.71	0.18
13. Number RT	PFNRTC	0.76	0.23	0.77	0.23
14. Combined RT	PFCRTC	1.03	0.30	1.10	0.31
15. Letter Coordination	PFLCOOR	3.44	7.86	3.47	7.86
16. Number Coordination	PFNCOOR	3.50	8.55	3.45	8.55
17. Combined Coordination	PFCCOOR	3.36	7.81	3.44	7.81
Shifting Attention Test					
18. Arrow Direction Thruput	SATADPUT	105.70	21.95	103.60	21.95
19. Discovery Rule Shifts	SATDIRUL	6.85	2.67	5.92	2.80
20. Discovery Rule Accuracy	SATDIACC	4.71	16.28	2.92	16.33
21. Discovery Failures	SATDIFAI	2.12	1.91	2.75	1.98
22. Discovery Perseverative Errors	SATDIPER	1.76	2.22	1.71	2.22
23. Instruction Thruput	SATINPUT	83.56	17.43	79.93	17.52
24. Arrow Color Thruput	SATACPUT	98.26	17.56	95.99	17.58
Dual Task					
25. Tracking "Alone" Error	DTTAABS	15.25	13.68	17.82	14.00
26. Tracking "Dual" Error	DTTDABS	45.16	26.82	52.22	27.56
27. Previous Number "Alone" Thruput	DTTPAPUT	153.23	186.40	152.37	186.70
28. Previous Number "Dual" Thruput	DTTPDPUT	121.49	157.44	102.53	157.96

Note. Means and standard deviations were corrected for range restriction using the multivariate method (Lawley, 1943).

The results of the CFA of Model 1 (Kay's [1995] 9-factor orthogonal model) indicated a poor fit. The CFI was .194 and the RMSEA was .207. This was not a surprising result for an orthogonal (uncorrelated) model as there is a long history of correlated abilities stretching back to the founding of factor analysis (Spearman, 1904).

Model 2, an oblique or correlated factors version of Model 1, fared considerably better. The CFI was .958 and the RMSEA was .049, both indicative of an acceptable fit. However, one of the correlations among the factors exceeded 1.0, probably a consequence of model mis-specification. It is apparent that this model does not fit the data thus rendering the interpretation of the fit indexes moot. Kay's (1995) 9-factor model, whether orthogonal or oblique, does not offer a good fit to the data.

Table 3. Correlation Matrix

Score	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	1000	161	028	044	062	-011	090	-013	-012	-002	033	-057	-059	-138	056	042	054	009	092	-014	-092	-019	076	049	-139	-121	-001	009
2	232	1000	158	291	-223	121	144	151	-045	-102	124	-099	-101	-277	-188	-175	-190	090	144	-012	-095	015	172	098	-139	-111	013	054
3	068	215	1000	371	-205	487	267	571	-280	-317	332	-345	-308	-369	-185	-176	-185	382	077	081	-011	-033	446	403	-090	-044	132	132
4	090	348	389	1000	-710	433	351	274	-247	-277	415	-330	-321	-426	-602	-574	-603	431	097	063	-072	-027	503	385	-262	-126	124	156
5	060	-197	-192	-689	1000	-339	-258	-159	117	172	-412	234	228	300	811	769	812	-349	-035	-094	034	021	-346	-244	299	095	-051	-081
6	017	157	491	442	-331	1000	336	321	-392	-247	420	-400	-356	-382	-285	-283	-283	507	-010	035	005	-025	559	515	-117	-089	207	215
7	134	221	293	381	-248	352	1000	212	-246	-200	291	-278	-216	-273	-199	-179	-198	371	012	009	-003	-083	404	370	-163	-128	128	130
8	018	195	583	293	-152	330	232	1000	-223	-275	242	-244	-211	-272	-133	-123	-132	236	032	008	-011	-044	310	272	-050	-002	095	048
9	-012	-034	-271	-237	112	-385	-238	-218	1000	479	-186	461	361	314	080	063	079	-566	028	-132	-017	-001	-555	-566	043	045	-184	-156
10	-019	-121	-324	-282	166	-248	-209	-282	476	1000	-184	287	251	272	138	128	139	-358	-023	-061	006	017	-370	-360	096	081	-104	-095
11	053	144	339	419	-406	422	299	249	-182	-189	1000	-268	-226	-291	-346	-326	-348	384	035	029	-015	-038	409	335	-209	-083	134	178
12	-074	-120	-351	-355	229	-402	-287	-252	458	290	-271	1000	784	528	193	212	195	-547	-026	-094	054	002	-549	-556	049	031	-176	-161
13	-083	-137	-321	-334	222	-363	-234	-224	357	257	-232	785	1000	463	188	288	193	-456	-029	-061	057	-004	-452	-455	047	056	-131	-136
14	-195	-370	-398	-457	277	-390	-320	-300	293	279	-298	523	471	1000	260	250	259	-407	-037	-066	033	068	-509	-472	140	109	-144	-139
15	047	-182	-180	-592	810	-281	-199	-131	075	136	-344	190	186	250	1000	934	996	-281	-077	-081	067	025	-311	-188	250	060	-046	-068
16	033	-171	-173	-565	767	-280	-180	-123	058	128	-325	210	286	243	933	1000	934	-258	-081	-080	071	012	-296	-173	222	057	-032	-066
17	045	-185	-181	-593	811	-280	-199	-131	074	137	-346	193	192	251	996	934	1000	-280	-077	-080	066	027	-308	-187	250	059	-044	-066
18	019	093	378	424	-345	503	369	234	-563	-355	384	-545	-455	-394	-278	-256	-278	1000	-016	109	023	-026	703	754	-115	-069	163	190
19	146	251	123	158	-036	027	080	069	029	-044	061	-047	-059	-124	-088	-092	-089	-003	1000	059	-802	009	-015	-039	-081	-119	012	035
20	-004	010	085	073	-095	042	022	011	-132	-064	035	-097	-065	-077	-085	-084	-084	111	075	1000	-038	009	097	094	-009	003	090	-008
21	-139	-192	-055	-127	035	-029	-062	-044	-017	024	-039	071	082	108	078	082	078	012	-817	-053	1000	-027	019	046	057	086	001	-022
22	-018	017	-030	-023	019	-022	-079	-044	-001	016	-035	001	-004	062	022	010	023	-024	012	011	-029	1000	-035	-016	042	029	002	049
23	102	205	451	508	-337	560	418	319	-544	-371	407	-550	-457	-514	-305	-291	-303	697	021	101	-012	-033	1000	747	-169	-113	219	251
24	065	117	403	386	-238	514	374	276	-561	-360	335	-556	-457	-466	-185	-171	-185	751	-015	097	024	-014	746	1000	-076	-070	180	193
25	-118	206	-123	-297	294	-141	-201	-075	043	109	-225	064	070	190	253	226	254	-123	-133	-021	105	038	-187	-089	1000	638	-028	-087
26	-159	-181	-079	-165	093	-114	-170	-028	044	092	-103	046	077	162	067	063	067	-080	-169	-010	131	024	-133	-083	653	1000	-004	-082
27	-002	004	121	115	-046	199	120	089	-181	-098	124	-171	-127	-127	-041	-027	-039	157	008	086	003	001	214	176	-022	002	1000	639
28	022	073	137	167	-082	221	139	053	-155	-097	180	-164	-141	-148	-071	-068	-069	191	052	-004	-038	049	256	197	-098	-093	637	1000

Notes. Names associated with scores 1-28 can be found in Table 2. Correlations above the diagonal are observed, while those below the diagonal were corrected for range restriction (Lawley, 1943). Decimal points were omitted.

Model 3 with six first-order factors had a CFI of .908 and a RMSEA of .072, again indicative of an acceptable fit. Models 2 and 3 cannot be tested directly against each other as one is not a proper subset of the other. Therefore, it is inappropriate to compare fit indices between these models. All correlations among the six first-order factors in Model 3 were within the bounds of ±1.0. We prefer this model for its interpretability and its parsimony. Table 4 shows the loadings of the scores on the factors and Table 5 shows the correlations among the first-order

Table 4. Factor Loadings for 6-Factor First-Order Model

	Factor					
Score	I	II	III	IV	V	VI
1. BDSACC						-.043
2. MATHPUT						.226
3. VSCPUT	.545					
4. SCDPUT	.349					.575
5. SDCIRACC						-.959
6. MTSPUT	.652					
7. MANPUT	.488					
8. DATSCPUT	.381					
9. DATIRTC		.682				
10. DATDRTC		.493				
11. ASCPUT	.483					
12. PFLRTC		.688			.085	
13. PFNRTC		.570			.103	
14. PFCRTC		.572			.162	
15. PFLCOOR					.998	
16. PFNCOOR					.935	
17. PFCCOOR					.998	
18. SATADPUT	.828		-.142			
19. SATDIRUL			.961			
20. SATDIACC			.075			
21. SATDIFAI			-.845			
22. SATDIPER			.015			
23. SATINPUT			-.121			
24. SATACPUT			-.155			
25. DTTAABS				1.188		
26. DTTDABS				.550		
27. DTTPAPUT	.244					
28. DTTPDPUT	.283					

Note. Factors I through VI are thruput, response time, shifting attention, tracking, pathfinder, and numeric.

factors. We interpreted the six lower order-factors by ordering the factor loadings of the tests from high to low and determining which tests had the greatest causal influence on the factor (Kim & Mueller, 1978). The first factor was clearly seen to be a measure of thruput; the second, response time; the third, selective attention; the fourth, tracking; the fifth, pathfinder; and the sixth, numerical operations.

Table 5. Correlations among First-Order Factors (6-Factor Model)

Factor	1	2	3	4	5	6
1. Thruput	1.0000					
2. Response Time	-0.9010	1.0000				
3. Selective Attention	0.1624	-0.0617	1.0000			
4. Tracking	-0.1584	0.0854	-0.0881	1.0000		
5. Pathfinder	-0.3400	0.1553	-0.0957	0.2274	1.0000	
6. Numerical Operations	0.4125	-0.2415	0.0577	-0.2752	-0.8452	1.0000

The correlations among the six first-order factors were then used to examine hierarchical models. There are two pairs of strongly correlated factors: thruput and response time and pathfinder and numerical operations. In keeping with previous findings, we tried a model with a single higher order factor and found it to be a poor fit to the data. Given past experience (Carretta & Ree, 1996, 1997; Carretta et al, in press; Ree & Carretta, 1994; Stauffer et al., 1996; Vernon, 1969), it is surprising that the single higher order factor model did not provide a better fit. As shown in Figure 1, the best fitting solution was for two higher order factors that were correlated. The correlation between these two higher order factors was -.316 indicating a shared source of variance and suggests a third level to the hierarchy. However, the existence of only two indicators (2 higher order factors) makes estimation of this third level factor difficult.

These two higher order factors are difficult to interpret. The first of these factors was derived from the thruput, response time, and shifting attention lower order factors. Thruput showed the highest loading on this factor. We have interpreted this factor as a measure of response efficiency. The other higher order factor included tracking, pathfinder, and numeric, with the latter two both having the largest loadings. We had difficulty interpreting this factor, but have labeled it procedural knowledge because of the requirement to apply multiple rules to complete the tests.

To understand the nature of the CogScreen-AE factors, we conducted a qualitative content analysis of the *tests* using Vernon's (1969) 3-level hierarchical model. Vernon's model has general cognitive ability (*g*) at the highest level and two broad major group factors called v:ed and k:m at the second level. V:ed is a verbal-educational factor and is typified by third-level minor group factors such as verbal and math. The k:m major group factor, called practical-mechanical is typified by minor group factors such as spatial and manual. This qualitative analysis suggested that only 2 of the 11 CogScreen-AE tests contribute to the v:ed factor while the remaining 9 are better categorized as measures of k:m.

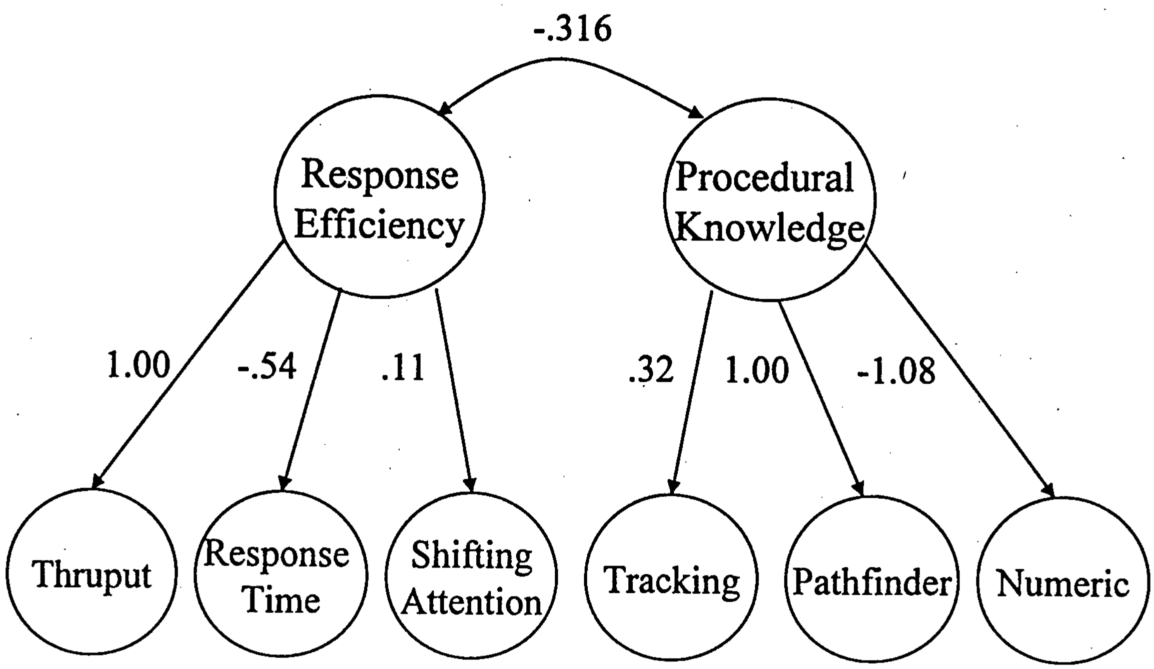


Figure 1. Factor Structure of the CogScreen-AE Test Battery

The higher order factor structure as determined by our factor analysis does not suggest a Vernon-like structure (Ree & Carretta, 1994; Carretta & Ree, 1997). The response efficiency factor contains tests that can be classified into both of Vernon's broad categories. The same is true of the procedural knowledge factor. These two higher order factors appear to include response modality and response process within each factor. CogScreen-AE offers potential for estimating pilot candidate performance based on these factors.

Kay (1995) offers interpretations of each test score (see his Tables 7 through 17). Because factors aggregate several tests, factors will be more reliable than individual test scores. This increased reliability raises the potential for finding validity for the CogScreen-AE. Carefully conducted validation studies are required to establish its validity and utility for use with pilots.

REFERENCES

Benson, A. J., & Gedye, J. L. (1963). *Logical process in the resolution of orientation conflict,* Report No. 259. Farnborough, UK: Royal Air Force, Institute of Aviation Medicine.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Lang (Eds.). *Testing statistical equation models* (pp. 136-162). Newbury Park, CA: Sage.

Callister, J. D., King, R. E., & Retzlaff, P. D. (1995). *Cognitive assessment of USAF pilot training candidates: Multidimensional Aptitude Battery and CogScreen-Aeromedical Edition* (AL/AO-TR-1995-0125). Brooks AFB, TX: Armstrong Laboratory, Aerospace Medical Directorate, Clinical Sciences Division.

Carretta, T. R., & Ree, M. J. (1993). Basic Attributes Test: Psychometric equating of a computer-based test. *The International Journal of Aviation Psychology, 3,* 189-201.

Carretta, T. R., & Ree, M. J. (1996). Factor structure of the Air Force Officer Qualifying Test: Analysis and comparison. *Military Psychology, 8,* 29-42.

Carretta, T. R., & Ree, M. J. (1997). Expanding the nexus of cognitive and psychomotor abilities. *International Journal of Selection and Assessment, 5,* 149-158.

Carretta, T. R., Retzlaff, P. D, Callister, J. D., & King, R. E. (1998). A comparison of two U. S. Air Force pilot aptitude tests. *Aviation, Space, and Environmental Medicine,69,* 931-935.

Hilton, T. F., & Dolgin, D. L. (1991). Pilot selection in the military of the free world. In R. Gal & A. D. Mangelsdorff (Eds.), *Handbook of Military Psychology* (pp. 81-101). NY: Wiley.

Imhoff, D. L., & Levine, J. H. (1981). *Perceptual-motor and cognitive performance task battery for pilot selection* (AFHRL-TR-80-27). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Jackson, D.N., Barton, C.F., & Blokker, H.C. (1992). *User's manual for the Multi-dimensional Aptitude Battery (MAB) software program.* London, Ontario, Canada: Sigma Assessment Systems.

Jensen, A. R. (1980). *Bias in mental testing*. NY: The Free Press.

Jöreskog, K., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International.

Jöreskog, K., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International.

Kay, G. G. (1995). *CogScreen-AE aeromedical edition professional manual.* Odessa, FL: Psychological Assessment Resources, Inc.

Kim, J., & Mueller, C. W. (1978). *Introduction to factor analysis: What it is and how to do it.* Beverly Hills, CA: Sage.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society of Edinburgh, Section A, 62, Part 1,* 28-30.

North, R. A., & Gopher, D. (1976). Measures of attention as predictors of flight performance. *Human Factors, 18,* 1-14.

Ree, M. J., & Carretta, T. R. (1994). Factor analysis of ASVAB: Confirming a Vernon-like model. *Educational and Psychological Measurement, 54,* 459-463.

Ree, M. J., & Carretta, T. R. (1996). Central role of *g* in military pilot selection. *The International Journal of Aviation Psychology, 6,* 111-123.

Ree, M. J., & Carretta, T. R. (1998). Computerized testing in the United States Air Force. *International Journal of Selection and Assessment, 6,* 82-89.

Ree, M. J., Carretta, T. R., Earles, J. A., & Albert, W. (1994). Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formulas. *Journal of Applied Psychology, 79,* 298-301.

Retzlaff, P. D., Callister, J. D., & King, R. E. (1996). *The computerized neuropsychological evaluation of US Air Force Pilots: Clinical procedures and data-based decision (*AL/AO-1996-0107). Brooks AFB, TX: Armstrong Laboratory, Aerospace Medical Directorate, Clinical Sciences Division.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analysis of Form , (*AFHRL-TR-86-68). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology, 15,* 201-293.

Spinner, B. (1991). Predicting success in primary flying school from the Canadian Automated Pilot Selection System: Derivation and cross-validation. *The International Journal of Aviation Psychology, 1,* 163-180.

Stauffer, J. M., Ree, M. J., & Carretta, T. R. (1996). Cognitive-components tests are not much more than g: An extension of Kyllonen's analyses. *The Journal of General Psychology, 123,* 193-205.

Thorndike, R. L. (1949). *Personnel selection.* NY: Wiley.

Vernon, P. E. (1969). *Intelligence and cultural environment.* London: Methuen.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* NY: The Psychological Corporation.